武铠蛱蝶线粒体基因组全序列测定和分析

王菊平[1]，宣善滨[1]，张育平[2]，杨　静[1]，曹天文[1,*]，马恩波[3]

(1. 山西省农业科学院植物保护研究所，太原 030031；2. 太原师范学院生物系，太原 030031；
3. 山西大学应用生物学研究所，太原 030006)

摘要：【目的】了解闪蛱蝶亚科属间及种间的分子系统进化关系。【方法】采用 PCR 步移法对武铠蛱蝶 *Chitoria ulupi* 线粒体基因组全序列进行了测定和分析。基于线粒体基因组 13 个蛋白质编码基因的核苷酸序列构建了 38 种鳞翅目昆虫的系统发育树。【结果】分析结果表明，武铠蛱蝶线粒体基因组全长 15 279 bp，包括 13 个蛋白质编码基因、22 个 tRNA 基因、2 个 rRNA 基因和一段长度为 391 bp 的 A + T 富含区，基因排列顺序与其他已知近缘种昆虫相同。武铠蛱蝶线粒体基因组中存在很高的 A + T 含量（79.9%）。13 个蛋白质编码基因中，*COII* 以 TTG 作为起始密码子，*COI* 以 CGA 作为起始密码子外，其余均为昆虫典型的起始密码子 ATN。*COII* 和 *ND*4 基因使用了不完全终止密码子 T，其余基因均以典型的 TAA 为终止密码子。在所测得的 22 个 tRNA 基因中，除 $tRNA^{Ser(AGN)}$ 缺少 DHU 臂外，其余 tRNA 均能形成典型的三叶草结构。与其他多数鳞翅目昆虫一样，武铠蛱蝶的 A + T 富含区中有一段由 ATAGAA 引导的保守的多聚 T 结构，长度为 21 bp，并散布着一些长短不一的串联重复单元。系统发育树结果显示，总科级别的系统发育关系为：卷蛾总科 +（凤蝶总科 +（螟蛾总科 +（夜蛾总科 + 蚕蛾总科 + 尺蛾总科）））；在蛱蝶科物种中，武铠蛱蝶与猫蛱蝶 *Timelaea maculate* 亲缘关系最近。【结论】基于分子标记构建的鳞翅目昆虫系统发育关系与传统的形态学分类结果基本一致。

关键词：鳞翅目；蛱蝶科；武铠蛱蝶；线粒体基因组；系统发育

中图分类号：Q969　　**文献标识码**：A　　**文章编号**：0454-6296(2015)03-0319-10

Sequencing and analysis of the complete mitochondrial genome of *Chitoria ulupi* (Lepidoptera: Nymphalidae)

WANG Ju-Ping[1], XUAN Shan-Bin[1], ZHANG Yu-Ping[2], YANG Jing[1], CAO Tian-Wen[1,*], MA En-Bo[3]

(1. Institute of Plant Protection, Shanxi Academy of Agricultural Sciences, Taiyuan 030031, China; 2. Biology department of Taiyuan Normal University, Taiyuan 030031, China; 3. Institute Applied Biology, Shanxi University, Taiyuan 030006, China)

Abstract: 【Aim】To explore the intergeneric and interspecific molecular and evolutionary relationships of Apaturinae. 【Methods】The complete mitochondrial genome of *Chitoria ulupi* was determined and analyzed by using polymerase chain reaction amplification and prime walking. Based on nucleotide sequences of 13 protein coding genes (PCGs), the phylogenetic tree of 38 lepidopteran species was constructed using maximum parsimony (MP) method. 【Results】The results indicated that the complete mitochondrial genome of *C. ulupi* is a circular molecule of 15 279 nucleotides, including 13 protein coding genes, 22 tRNA genes, 2 rRNA genes and a 391 bp A + T-rich region. The arrangement of genes in *C. ulupi* is consistent with the sequence of other closely related species. The mitochondrial genome of *C. ulupi* is biased toward a high A + T content (79.9%). All protein coding genes start with a typical ATN initiation codon, except that the *COI* starts with the CGA codon and the *COII* begins with the TTG codon. Majority of the 13 PCGs in *C. ulupi* have a complete termination codon (TAA or TAG), except for the *COII* and *ND*4 genes which have incomplete stop codons (T). All tRNAs show the classic clover-leaf structure, except that the dihydrouridine (DHU) arm of $tRNA^{Ser(AGN)}$ forms a simple loop. The *C. ulupi* A + T-rich region contains some conserved structures including a structure combing the motif ATAGAA, 21 bp poly(T) stretch and a microsatellite $(AT)_n$ element, which is similar to those

基金项目：国家自然科学基金项目(40871034)；山西省科技基础条件平台建设项目(2013091004-0104)；山西省自然科学基金项目(2011011033-2)

作者简介：王菊平，女，1976 年 11 月生，山西长治人，助理研究员，研究方向为昆虫分子系统学，E-mail：jp76@163.com

* 通讯作者 Corresponding author, E-mail：ctwen@126.com

收稿日期 Received：2014-12-24；接受日期 Accepted：2015-02-12

found in the mitogenomes of other related species. The results of phylogenetic tree showed that the phylogenetic relationship at the superfamily level is Tortricoidea + (Papilionoidea + (Pyraloidea + (Noctuoidea + Bombycoidea + Geometroidea))), and within the Nymphalidae group, *C. ulupi* is closely related with *Timelaea maculate*. 【Conclusion】 The phylogenetic relationship of lepidopteran species based on molecular markers is basically consistent with the results of the traditional morphological classification.

Key words: Lepidoptera; Nymphalidae; *Chitoria ulupi*; mitochondrial genome; phylogeny

动物线粒体基因组具有相对较小(约为 15 ~ 19 kb)、结构简单、母系遗传和进化速率快等特点(黄原, 1998)。在分子生物学研究方面是研究 DNA 结构与复制转录的良好模型。此外,线粒体基因组作为一种分子标记,包含多个进化速率各异的基因,广泛应用于动物界不同等级阶元系统发育的研究(Ballard, 2000; Buckley *et al*., 2001; Cameron *et al*., 2007)。

武铠蛱蝶 *Chitoria ulupi* 隶属于鳞翅目(Lepidoptera),蛱蝶科(Nymphalidae),闪蛱蝶亚科(Apaturinae)。分布于我国辽宁、河北、山西、浙江、福建、江西、四川、云南等地。幼虫以榆科的珊瑚朴 *Celtis julianae* 为寄主。目前国内对武铠蛱蝶的研究仅有成虫形态及生境等方面的报道(周尧等, 1994, 1998),尚无对武铠蛱蝶线粒体基因组全序列的研究报道。

闪蛱蝶亚科是昆虫纲鳞翅目蛱蝶科的一个重要类群,包括许多农林业重要害虫。该亚科物种对环境的敏感性强,地理种群分化较多,亚科的构成、属间和种间系统关系一直未达成共识。本研究以武铠蛱蝶作为研究对象,对其线粒体基因组全序列进行了测定和分析,并与相关的近缘物种线粒体基因组进行了比较分析,以期为蛱蝶科系统学和分类学的研究提供分子生物学的证据。

1 材料与方法

1.1 标本的采集和保存

武铠蛱蝶 2009 年 7 月 12 日采自山西省中条山(34.87°N,110.86°E,海拔 910 m),采集到的昆虫活体保存在 -80℃冰箱中备用。

1.2 基因组 DNA 提取

挑选保存完好的蝴蝶标本,取其胸部肌肉,采用 OMEGA Insect DNA Kit 试剂盒结合酚氯仿法提取样品总 DNA,抽提完毕后,用 1.0% 琼脂糖凝胶电泳检测其纯度及浓度,并于 -20℃保存备用。

1.3 引物设计及 PCR 扩增

参考 Simon 2006 年发表的线粒体通用引物序列和目前已发表的鳞翅目相近物种引物(Hwang *et al*., 2001; Hong *et al*., 2009)进行首轮扩增。其他未能得到的序列通过在 GenBenk 中下载相关昆虫部分线粒体基因序列进行 ClustalX(Thompson *et al*., 1997)对比寻找相关保守区,根据引物设计的基本原则(张新宇和高燕宁, 2004),利用引物设计软件 Primer Primier(Ver. 5.0)(Singh *et al*., 1998)设计相关引物。本实验共用引物 22 对,覆盖了武铠蛱蝶线粒体基因组所有序列。

按照引物序列,采用 TaKaRa La Taq 聚合酶扩增各个片段,所有 PCR 反应均在冰上操作。PCR 反应体系为 25 μL,其中 10 × LA PCR Buffer 2.5 μL, dNTPs 4 μL(2.5 mmol/L),上下游引物各 1 μL(10 μmol /L),模板 DNA 1 μL,TaKaRa LA Taq 酶 0.25 μL(5 U/μL),灭菌水补足至 25 μL。少数引物在进行 PCR 扩增反应时对模板 DNA 和引物的量作适当调整。PCR 扩增条件为:95℃预变性 5 min;95℃变性 50 s,48℃复性 50 s,68℃延伸 2 min,30 个循环;68℃延伸 10 min。采用 1% 的琼脂糖凝胶电泳检测 PCR 产物。

1.4 序列测定

目的 PCR 扩增产物通过直接测序法获得目的片段信息。所有片段均为双向测通,未能测通的序列则根据自测序列再设计引物,PCR 扩增双向测序后补全线粒体基因组全部序列。

1.5 序列的拼接与注释

得到所有序列片段测序结果后,利用 Sequencher 4.8(Gene Codes Corporation, 2008)软件对所有序列进行拼接。线粒体全基因组各基因的定位及注释参考鳞翅目已测得的鳞翅目昆虫全线粒体基因组注释结果(徐浪等, 2007; Kim *et al*., 2009; Jiang *et al*., 2009)进行,通过 MEGA4.0(Tamura *et al*., 2007)比对以及 Blast 在线搜索分析相结合来依次注释 13 个蛋白质编码基因(protein coding genes,

PCGs）、22 个 tRNA 基因、2 个 rRNA 基因和 AT 丰富区的位置。tRNA 基因二级结构的预测主要是通过 tRNAscan-SE 1.21 在线预测并进行人工校正。

1.6　系统发育分析

基于全线粒体基因组的 13 个蛋白质编码基因核苷酸序列建立 MP 树，研究中涉及到的鳞翅目昆虫 38 种，其中蛱蝶科昆虫 12 种，并选择了与鳞翅目关系较近的果蝇 *Drosophila yakuba*、冈比亚按蚊 *Anopheles gambiae*、意大利蜜蜂 *Apis mellifera ligustica* 和小峰熊蜂 *Bombus hypocrita* 作为外群。将 38 种鳞翅目昆虫和 4 个外群线粒体基因组中的 13 个蛋白质编码基因序列串联后，通过 Clustalx 1.83 软件进行序列比对，避免移码。比对之后采用最大简约法（maximum parsimony method，MP）进行分析，MP 法所用软件为 PAUP 4.0，其参数如下：heuristic search，TBR branch swapping。用 Bootstrap（1 000 replications）来检验各节点的置信度，使用的参数与搜索最简约树相同。在搜索最简约树时，所有位点被认为碱基处理作 Fitch 性状（无序）。是具有 4 种状态（A，T，C，G）的无序位点，排序中的裂缝作为缺失处理。

2　结果

2.1　武铠蛱蝶 mtDNA 基因组结构

武铠蛱蝶的线粒体基因组全长为 15 279 bp（全序列已经提交到 GenBank，登录号：KP284554，基因组成与典型的后生动物线粒体 DNA 相似（Wolstenholme，1992），包含 13 个蛋白质编码基因（PCGs）、2 个 rRNA 基因（16S *rrnL* 和 12S *rrnS*）、22 个 tRNA 基因及一个 AT 富含区（A + T-rich region）。在线粒体基因组编码的 37 个基因中，N 链（minority-strand）编码了 14 个基因，包括 4 个蛋白质编码基因，8 个 tRNA 基因及 2 个 rRNA 基因；其他 23 个基因由 J 链编码。*rrnS* 基因和 $tRNA^{Met}$ 基因之间 391 bp 确定为 AT 富含区。基因排列顺序与其近缘种昆虫的基因排列顺序一致（Zhang *et al.*，2012；Chen *et al.*，2012；Wang *et al.*，2013）。其中 *trnM-trnI-trnQ* 的排列方式与假想昆虫祖先的排列方式 *trnI-trnQ-trnM* 不同（Boore *et al.*，1998），但这种排列方式在鳞翅目昆虫中却非常典型。

武铠蛱蝶线粒体基因组中有 16 处基因重叠，共 111 bp，最长一处位于 *ND5* 和 $tRNA^{Phe}$ 之间，重叠序列为 56 bp；基因间隔区有 12 处，共 168 bp，$tRNA^{Thr}$ 和 *ND4L* 之间间隔最长为 59 bp，其次是 $tRNA^{Gln}$ 和 *ND2* 之间，间隔 52 bp。既无间隔又无重叠的区域共有 9 处（表 1）。武铠蛱蝶的基因结构及基因顺序见图 1 及表 1。

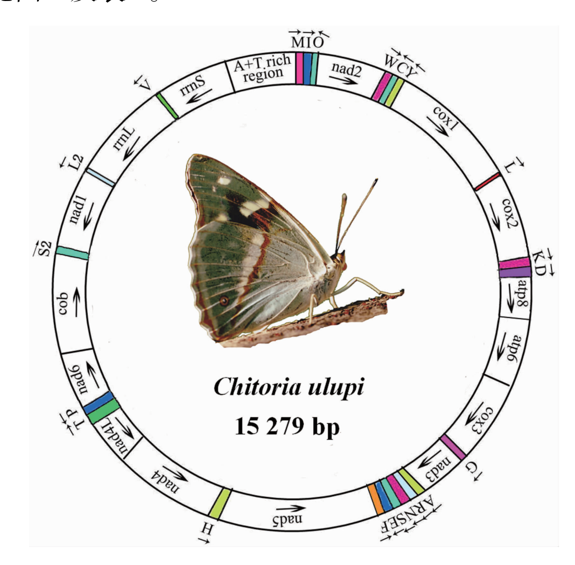


图 1　武铠蛱蝶线粒体基因组结构图

Fig. 1　Gene map of the mitogenome of *Chitoria ulupi*

2.2　武铠蛱蝶线粒体基因组核苷酸组成

武铠蛱蝶线粒体基因组中 A，T，C 和 G 碱基含量分别为 38.9%，41.0%，12.4% 和 7.8%，T 含量最高而 G 含量最低。A + T 含量为 79.9%，G + C 含量为 20.2%。蛋白质编码基因总共含 3 739 个密码子，不包含终止密码子，A + T 含量为 78.2%，比全序列的 A + T 含量稍低。蛋白质编码基因密码子第 1 位点的 AT 偏向性（80.4%）高于第 2 位点（76.5%）和第 3 位点（77.7%）。22 个 tRNA 基因 A + T 含量为81.1%，*rrnL* 基因的 A + T 含量为 84.3%，*rrnS* 基因的 A + T 含量为 85.1%，控制区的 A + T 含量最高为 90.3%（表 2）。

武铠蛱蝶线粒体基因组中存在很高的 A + T 含量，这与其他鳞翅目昆虫是一样的。武铠蛱蝶线粒体基因组中 A + T 含量与已知的几种近缘种相比是较低的。武铠蛱蝶 *C. ulupi*（79.9%）= 大紫蛱蝶 *Sasakia charonda*（79.9%）< 柳紫闪蛱蝶 *Apatura ilia*（80.5%）= 细带闪蛱蝶 *Apatura metis*（80.5%）< 猫蛱蝶 *Timelaea maculate*（81.1%）< 黑紫蛱蝶 *Sasakia funebris*（81.2%）（表 3）（王菊平等，2012；Zhang *et al.*，2012；Chen *et al.*，2012；Wang *et al.*，2013；曹天文等，2013）。分析结果符合昆虫线粒体基因组中 A + T 含量偏向性的特征（Tzeng *et al.*，1992；Lee and Kocher，1995；Noack *et al.*，1996）。

表 1 武铠蛱蝶线粒体基因组注释

Table 1 Organization of the mitochondrial genome of *Chitoria ulupi*

基因 Gene	方向 Direction	位置 Location	大小 Size	反密码子 Anticodon	起始密码子 Start codon	终止密码子 Stop codon	基因间的间隔* Intergenic length*
tRNA^{Met}	F	1 – 68	68	CAT			
tRNA^{Ile}	F	69 – 133	65	GAT			0
tRNA^{Gln}	R	131 – 199	69	TTG			–3
*ND*2	F	252 – 1 265	1 014		ATT	TAA	52
tRNA^{Trp}	F	1 264 – 1 331	68	TCA			–2
tRNA^{Cys}	R	1 324 – 1 385	62	GCA			–8
tRNA^{Tyr}	R	1 392 – 1 456	65	GTA			6
COI	F	1 467 – 3 002	1 536		CGA	TAA	10
tRNA$^{Leu(UUR)}$	F	2 998 – 3 065	68	TAA			–5
COII	F	3 066 – 3 744	679		TTG	T – tRNA	0
tRNA^{Lys}	F	3 742 – 3 812	71	CTT			–3
tRNA^{Asp}	F	3 812 – 3 877	66	GTC			–1
*ATP*8	F	3 878 – 4 042	165		ATC	TAA	0
*ATP*6	F	4 036 – 4 713	678		ATG	TAA	–7
COIII	F	4 713 – 5 501	789		ATG	TAA	–1
tRNA^{Gly}	F	5 504 – 5 569	66	TCC			2
*ND*3	F	5 567 – 5 923	357		ATA	TAA	–3
tRNA^{Ala}	F	5 931 – 5 995	65	TGC			7
tRNA^{Arg}	F	5 995 – 6 058	64	TCG			–1
tRNA^{Asn}	F	6 065 – 6 130	66	GTT			6
tRNA$^{Ser(AGN)}$	F	6 129 – 6 189	61	GCT			–2
tRNA^{Glu}	F	6 192 – 6 256	65	TTC			2
tRNA^{Phe}	R	6 255 – 6 321	67	GAA			–2
*ND*5	R	6 266 – 8 056	1 791		ATC	TAA	–56
tRNA^{His}	R	8 054 – 8 121	68	GTG			–3
*ND*4	R	8 122 – 9 460	1 339		ATG	T – tRNA	0
ND4L	R	9 460 – 9 744	285		ATG	TAA	–1
tRNA^{Thr}	F	9 804 – 9 868	65	TGT			59
tRNA^{Pro}	R	9 869 – 9 932	64	TGG			0
*ND*6	F	9 935 – 10 462	528		ATT	TAA	2
CytB	F	10 467 – 11 618	1 152		ATG	TAA	4
tRNA$^{Ser(UCN)}$	F	11 617 – 11 683	67	TGA			–2
*ND*1	R	11 701 – 12 639	939		ATG	TAA	17
tRNA$^{Leu(CUN)}$	R	12 641 – 12 714	74	TAG			1
rrnL	R	12 715 – 14 048	1 334				0
tRNA^{Val}	R	14 049 – 14 112	64	TAC			0
rrnS	R	14 113 – 14 888	776				0
控制区 Control region		14 889 – 15 279	391				0

*基因间的间隔一列中，正数表示基因间隔碱基数，负数表示基因重叠碱基数。*In the column of intergenic length the positive number indicates interval base pairs between genes while the overlapping base between genes.

表 2　武铠蛱蝶线粒体基因组核苷酸组成

Table 2　Nucleotide composition of the mitochondrial genome of *Chitoria ulupi*

	大小 Size（bp）	A%	T%	G%	C%	（A＋T）%	（G＋C）%
全基因组 mtDNA	15 279	38.9	41.0	7.8	12.4	79.9	20.2
蛋白质编码基因 PCGs	11 252	33.2	45.0	11.2	10.6	78.2	21.8
密码子第 1 位点 1st codon site	3 739	34.9	45.5	10.1	9.4	80.4	19.5
密码子第 2 位点 2nd codon site	3 739	33.7	42.8	12.2	11.3	76.5	23.5
密码子第 3 位点 3rd codon site	3 739	31.0	46.7	11.1	11.1	77.7	22.2
tRNA	1 469	41.5	39.6	10.8	8.2	81.1	19.0
rrnL	1 334	45.5	38.8	10.6	5.0	84.3	15.6
rrnS	776	44.6	40.5	10.1	4.9	85.1	15.0
控制区 Control region	391	42.7	47.6	3.8	5.9	90.3	9.7

表 3　已知的闪蛱蝶亚科 6 种近缘种线粒体基因组的碱基组成

Table 3　Mitogenomic nucleotide compositions of the 6 closely related species of the Apaturinae

物种 Species	mtDNA		密码子个数[a] Number of codons[a]	蛋白质编码基因 A＋T 含量[b] A＋T content（%）in PCG[b]	*rrnL*		*rrnS*		A＋T 丰富区 A＋T-rich region		GenBank 登录号 GenBank accession no.
	大小 Size（bp）	（A＋T）%			大小 Size（bp）	（A＋T）%	大小 Size（bp）	（A＋T）%	大小 Size（bp）	（A＋T）%	
Chitoria ulupi	15 279	79.9	3 739	78.2	1 334	84.3	776	85.1	391	90.3	KP284554
Sasakia charonda	15 233	79.9	3 732	78.4	1 330	84.0	776	84.9	381	91.6	JX119051
Apatura ilia	15 242	80.5	3 711	78.9	1 333	86.0	776	84.9	403	92.5	NC_016062
Apatura metis	15 236	80.5	3 717	78.9	1 333	84.5	779	84.8	394	92.9	NC_015537
Timelaea maculata	15 178	81.1	3 734	79.7	1 332	85.0	777	85.9	382	93.2	NC_021090
Sasakia funebris	15 233	81.2	3 736	80.0	1 333	84.6	772	85.5	370	93.0	NC_022134

[a]在总的密码子计算中没有包括终止密码子 Termination codons were excluded in total codon count；[b]蛋白质编码基因 Protein coding genes.

2.3　武铠蛱蝶线粒体基因组蛋白质编码基因

武铠蛱蝶线粒体基因组含 13 个蛋白质编码基因，其中 *ND*2 和 *ND*6 以 ATT 作为起始密码子，*ATP*6，*COIII*，*ND*4，*ND*4*L*，*CytB* 和 *ND*1 以 ATG 作为起始密码子，*ATP*8 和 *ND*5 以 ATC 作为起始密码子，*ND*3 以 ATA 作为起始密码子，*COII* 以 TTG 作为起始密码子，*COI* 以 CGA 作为起始密码子。而在其他 5 种近缘种中，*COII* 均以 ATG 作为起始密码子，*ATP*8 和 *ND*5 均以 ATT 作为起始密码子；柳紫闪蛱蝶 *COIII* 以 ATA 作为起始密码子；细带闪蛱蝶 *ND*3 以 ATT 作为起始密码子；黑紫蛱蝶 *ND4L* 以 ATA 作为起始密码子；大紫蛱蝶、柳紫闪蛱蝶、细带闪蛱蝶和黑紫蛱蝶的 *ND*6 以 ATA 作为起始密码子，猫蛱蝶则以 ATC 作为起始密码子（表 4）。

终止密码子有以 TAA 或 TAG 标准密码子作为终止密码子，也有以单独的 T 作为终止密码子的蛋白质编码基因，*COII* 和 *ND*4 以单独的 T 作为终止密码子，*ND*2，*COI*，*ATP*8，*ATP*6，*COIII*，*ND*3，*ND*5，*ND4L*，*ND*6，*CytB* 和 *ND*1 以 TAA 作为终止密码子（表 4）。

全部蛋白质编码基因的密码子和氨基酸使用情况的统计分析显示，使用最频繁的氨基酸为 Leu，Ile，Phe 和 Met，使用最频繁的密码子为 TTA，ATT，TTT 和 ATA，反映了核苷酸组成的 AT 偏好性。对于检测到的 tRNA 基因，使用最频繁的氨基酸并不总是与 tRNA 的反密码子对应的。

2.4　武铠蛱蝶线粒体基因组 tRNA 基因和 rRNA 基因

根据 tRNAScan-SE 预测结果，并辅助人工校正的方法预测出武铠蛱蝶线粒体的 22 个 tRNA 结构。武铠蛱蝶 22 个线粒体 tRNA 中，除 $tRNA^{Leu}$ 和 $tRNA^{Ser}$ 所对应的 tRNA 有 2 个以外，其他的氨基酸都只有 1 个 tRNA 与之对应。除了 $tRNA^{Ser(AGN)}$ 缺少 DHU 臂外，其余 21 个 tRNA 均可形成典型的三叶草结构（图 2）。

武铠蛱蝶线粒体基因组中 22 个 tRNA 基因的排列顺序与其近缘种昆虫的相同，基因长度在 61～79 bp 之间。每种 tRNA 的反密码子都与已知鳞翅目蛱蝶科物种（*E*. *autonoe* 和 *A*. *issoria*）的反密码子相同（Hu *et al*., 2010；Kim *et al*., 2010）。tRNA 的二级结构也与已知蛱蝶科物种非常类似，这种现象与许多后生动物相一致（张方和米志勇，1998；Hong

表 4 闪蛱蝶亚科 6 种近缘种的起始密码子和终止密码子的使用情况

Table 4 Start and stop codon usages of the 6 closely related species of the Apaturinae

基因 Gene	*Chitoria ulupi*		*Sasakia charonda*		*Sasakia funebris*		*Apatura ilia*		*Apatura metis*		*Timelaea maculata*	
	起始密码子 Start codon	终止密码子 Stop codon	起始密码子 Start codon	终止密码子 Stop codon	起始密码子 Start codon	终止密码子 Stop codon	起始密码子 Start codon	终止密码子 Stop codon	起始密码子 Start codon	终止密码子 Stop codon	起始密码子 Start codon	终止密码子 Stop codon
*ND*2	ATT	TAA	ATT	TAA	ATT	TAA	ATT	TAA	ATT	TAA	ATT	TAA
COI	CGA	TAA	CGA	TAA	CGA	TAA	CGA	TAA	CGA	TAA	CGA	TAA
COII	TTG	T	ATG	T	ATG	T	ATG	T	ATG	T	ATG	T
*ATP*8	ATC	TAA	ATT	TAA	ATT	TAA	ATT	TAA	ATT	TAA	ATT	TAA
*ATP*6	ATG	TAA	ATG	TAA	ATG	TAA	ATG	TAG	ATG	TAA	ATG	TAA
COIII	ATG	TAA	ATG	TAA	ATG	TAA	ATA	TAA	ATG	TAA	ATG	TAA
*ND*3	ATA	TAA	ATA	TAG	ATA	TAA	ATA	T	ATT	TAG	ATA	TAG
*ND*5	ATC	TAA	ATT	TAA	ATT	TAA	ATT	T	ATT	T	ATT	TAA
*ND*4	ATG	T	ATG	T	ATG	T	ATG	T	ATG	T	ATG	T
ND4L	ATG	TAA	ATG	TAA	ATA	TAA	ATG	TAG	ATG	TAA	ATG	TAA
*ND*6	ATT	TAA	ATA	TAA	ATA	TAA	ATA	TAA	ATA	TAA	ATC	TAA
CytB	ATG	TAA	ATG	TAA	ATG	TAA	ATG	TAA	ATG	TAA	ATG	TAA
*ND*1	ATG	TAA	ATG	TAA	ATG	TAG	ATG	TAA	ATG	TAA	ATG	TAA

et al., 2009; Liao *et al*., 2010)。

在所预测的 22 种 tRNA 二级结构中总共出现了 21 对碱基错配,其中有 20 对为 GU 错配,其余 1 对为 AA 错配。AA 错配发生在 $tRNA^{His}$ 的 T + C 臂上(图 2)。大紫蛱蝶、黑紫蛱蝶、猫蛱蝶、细带闪蛱蝶和柳紫闪蛱蝶的 tRNA 二级结构均含有 G = U, U = U 非经典配对,这种错配现象在鳞翅目之外的许多昆虫类群中也有发现(Bae *et al*., 2004)。

rRNA 基因是线粒体基因组中进化最慢、最为保守的基因,仅根据核苷酸序列无法精确定位 rRNA 基因的起始位置,因此通常被认为伸达两侧基因的边界(Boore, 2006)。*rrnS* 在基因组中的位置位于 $tRNA^{Val}$ 和控制区之间, *rrnL* 位于 $tRNA^{Leu(CUN)}$ 和 $tRNA^{Val}$ 之间。武铠蛱蝶的 *rrnL* 和 *rrnS* 长度分别是 1 334 bp 和 776 bp, AT 含量分别为 84.3% 和 85.1%。武铠蛱蝶的 *rrnL* 长度与几个近缘种相比差别均较小,如:黑紫蛱蝶 1 333 bp, 细带闪蛱蝶 1 333 bp, 柳紫闪蛱蝶 1 333 bp, 猫蛱蝶 1 332 bp 和大紫蛱蝶 1 330 bp,相差 1 ~4 bp。*rrnS* 长度差别也不大,依次为细带闪蛱蝶 779 bp,猫蛱蝶 777 bp,大紫蛱蝶和柳紫闪蛱 776 bp,黑紫蛱蝶 772 bp,相差 2 ~7 bp。(表 3)。

2.5 武铠蛱蝶线粒体基因组 A + T 富含区

线粒体基因组 AT 富含区为线粒体基因组主要的非编码区,武铠蛱蝶 A + T 富含区位于 *rrnS* 基因和 $tRNA^{Met}$ 基因之间,长度为 391 bp, A + T 含量为 90.3%。几个近缘物种 A + T 富含区长度相差较大,依次为:柳紫闪蛱蝶(403 bp) > 细带闪蛱蝶(394 bp) > 武铠蛱蝶 > 猫蛱蝶(382 bp) > 大紫蛱蝶(381 bp) > 黑紫蛱蝶(370 bp),相差 11 ~33bp; A + T 含量则为几个近缘种最低的,依次为猫蛱蝶(93.2%) > 黑紫蛱蝶(93.0%) > 细带闪蛱蝶(92.9%) > 大紫蛱蝶(91.6%) > 武铠蛱蝶(表 3)。

在 *rrnS* 基因的下游有一段由 ATAGAA 引导的保守的长度为 21 bp 的多聚 T 结构。此外,在武铠蛱蝶线粒体 A + T 富含区中存在 2 段长度为 35 bp 的重复单元,其重复序列之间的间隔为 0。并且与其他鳞翅目昆虫比较也存在$(AT)_8$ 结构。在重复序列上游存在 10 bp 的多聚 A 结构(图 3)。

2.6 武铠蛱蝶系统发育关系

以果蝇 *D. yakuba*、冈比亚按蚊 *A. gambiae*、意大利蜜蜂 *A. mellifera ligustica* 和小峰熊蜂 *B. hypocrita* 作为外群, 基于 38 种鳞翅目物种全线粒体基因组的 13 个蛋白质编码基因核苷酸序列建立 MP 树。结果显示(图 4)总科级别的系统发育关系为:卷蛾总科 Tortricoidea + (凤蝶总科 Papilionoidea + (螟蛾总科 Pyraloidea + (夜蛾总科 Noctuoidea + 蚕蛾总科 Bombycoidea + 尺蛾总科 Geometroidea)))。

所分析的 38 种鳞翅目昆虫基本分为 3 个大的聚类簇:卷蛾总科卷叶蛾科(Tortricidae)的 3 个种类聚在一起,构成第 1 聚类簇;凤蝶总科的粉蝶科

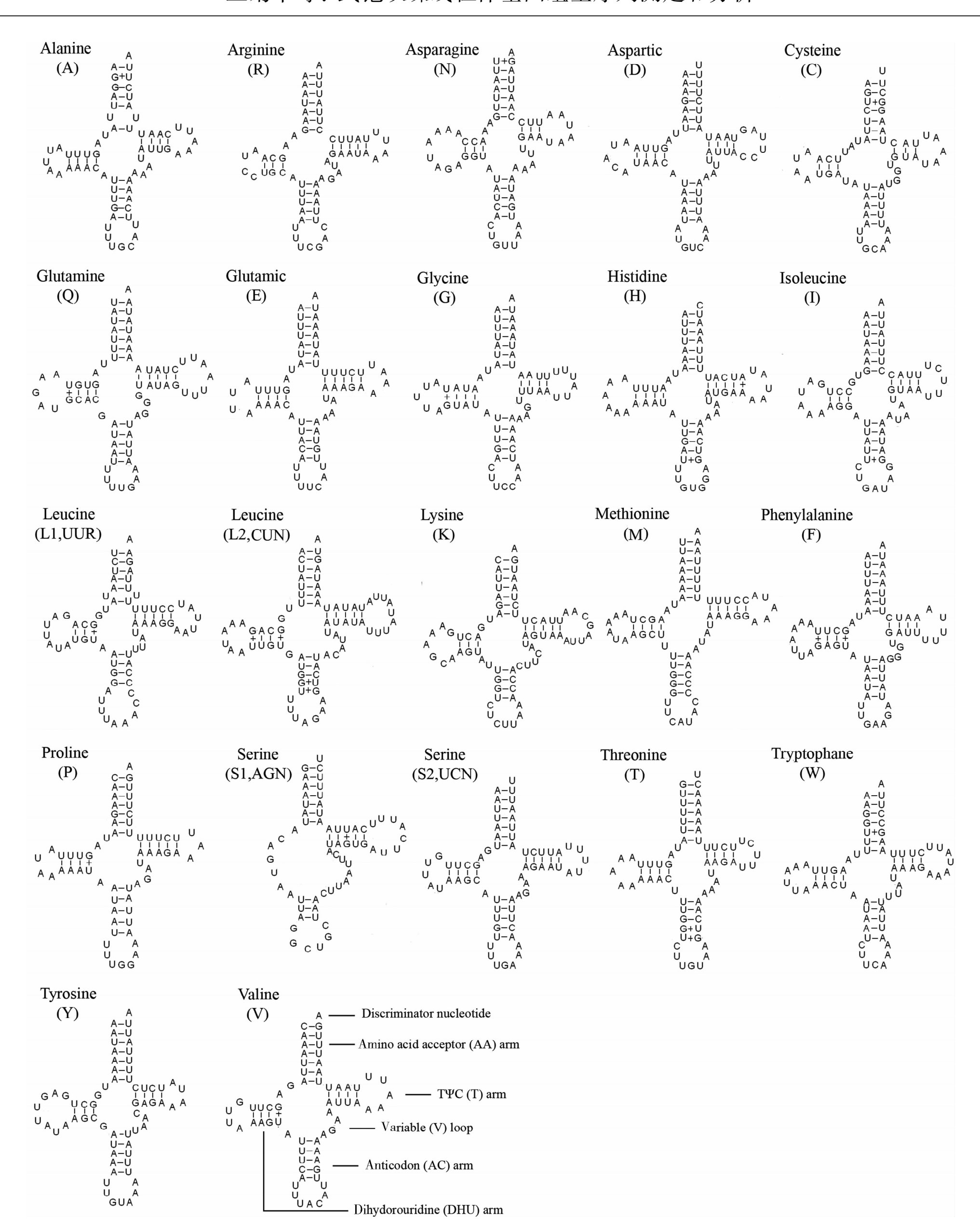


图 2　武铠蛱蝶 tRNA 的二级结构

Fig. 2　Secondary structures for tRNA genes from *Chitoria ulupi* mtDNA

rrnS-14 888 bp

ATTATCACTATTTTTCACATAGATTTTTTTTTTTTTTTTTTTTTATATTAAATGTATAATA

轻链复制起点

Origin of light strand replication

GACATTAATAAATTTTTATTATTTATTCTTTCTCTTTTTTCATAATATTGTTATTAAAAA

TTAATATCAATATAAGTTGTATATTAATCCAATTAAATAACAATAAATTAATAAAATTA

TAATTAATTTAAAAAAAAAATTAATTAATTATTAATATATTAATTAATTAAATATTTAAT

多聚 A 结构　　**35 bp 的重复单元**

Poly-A element　　**35 bp repeat element**

TAATTATTAATATATTAATTAATTAAATATTTAATATATATATATATATATACATGTGTGTGC

35 bp 的重复单元　　**微卫星(AT)n 单元**

35 bp repeat element　　**Macrosatellite (AT)n element**

GTATGCATACATATATGTACATTAAATTAACAAAAACTATTTATTTATATAGTTTATTAA

ACCATTTTTAATAATTTTTACATAAAAGAA　15 279 bp-- tRNA[Met]

图 3　武铠蛱蝶 AT 富含区结构

Fig. 3　Structure of the A + T-rich region of the mitochondrial genome of *Chitoria ulupi*

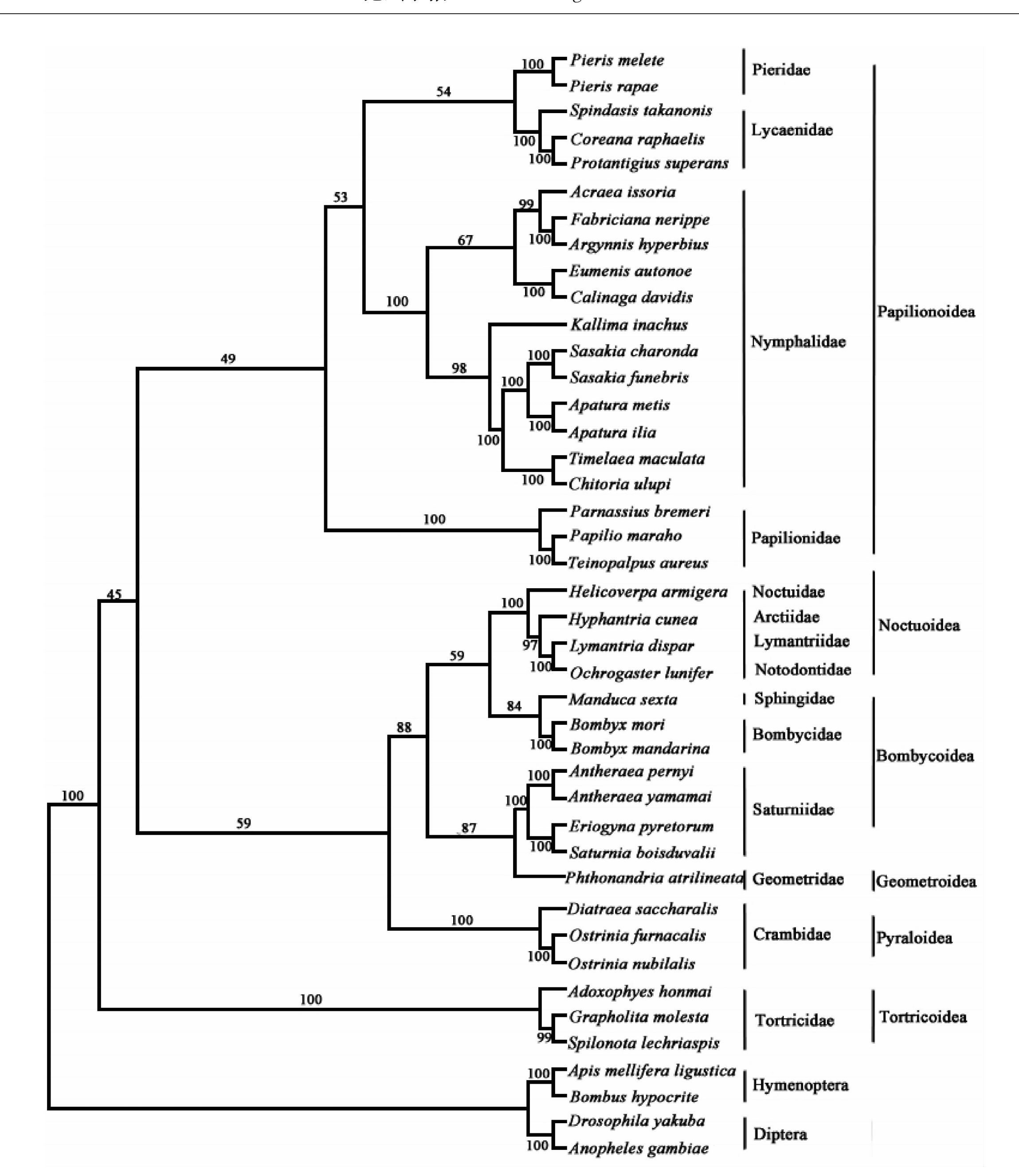


图 4　基于 38 种鳞翅目昆虫的 13 个蛋白质编码基因核苷酸序列的最大简约树

Fig. 4　Maximum parsimonious phylogenetic tree of 38 lepidopteran species based on the nucleotides of 13 protein coding genes

分支节点的数值表示自引导值。Bootstrap support values are indicated at each node.

(Pieridae)与灰蝶科(Lycaenidae)的种类先聚,再与蛱蝶科(Nymphalidae)的种类相聚,最后与凤蝶科(Papilionidae)的种类聚在一起,构成第 2 聚类簇;蚕蛾总科的天蛾科(Sphingidae)、蚕蛾科(Bombycidae)、天蚕蛾科(Saturniidae)和尺蛾总科的尺蛾科(Geometridae)的种类先聚在一起,再与夜蛾总科夜蛾科(Noctuidae)、灯蛾科(Arctiidae)、毒蛾科(Lymantriidae)和舟蛾科(Notodontidae)的种类相聚,最后与螟蛾总科的草螟科(Crambidae)的种类聚在一起,形成第 3 聚类簇。总科内各科及科内各物种优先以很高的置信度相聚,然后再与其他科相聚。

蛱蝶科 12 个物种又分为 2 个小簇,第 1 簇:(苎麻珍蝶 *Acraea issoria* +(蟾福蛱蝶 *Fabriciana nerippe* +斐豹蛱蝶 *Argynnis hyperbius*))+(仁眼蝶 *Eumenis autonoe* 和大卫绢蛱蝶 *Calinaga davidis*);第 2 簇为:(((大紫蛱蝶 *Sasakia charonda* +黑紫蛱蝶 *Sasakia funebris*)+(细带闪蛱蝶 *Apatura metis* +柳紫闪蛱蝶 *Apatura ilia*))+(猫蛱蝶 *Timelaea maculata* +武铠蛱蝶 *Chitoria ulupi*))+枯叶蛱蝶 *Kallima inachus*。

3 讨论

在武铠蛱蝶线粒体基因组中，在 *COI* 基因起始区及紧邻前面 tRNA[Tyr] 基因的区域都没有发现标准的 ATN 起始密码子，这种现象在其他近缘种中同样存在。我们将武铠蛱蝶和所有目前已知蝶类 *COI* 基因起始区附近的序列进行了比对，参照其近缘种推定 CGA 作为 *COI* 的起始密码子。武铠蛱蝶线粒体基因组中的 *COII* 以 TTG 作为起始密码子，*ATP*8 和 *ND*5 以 ATC 作为起始密码子。而在其他 5 种近缘种中，*COII* 均以 ATG 作为起始密码子，*ATP*8 和 *ND*5 均以 ATT 作为起始密码子。在终止密码子上，武铠蛱蝶与其他 5 种近缘种均有不同之处：大紫蛱蝶和猫蛱蝶 *ND*3 以 TAG 作为终止密码子；黑紫蛱蝶 *ND*1 以 TAG 为终止密码子；细带闪蛱蝶 *ND*3 以 TAG，*ND*5 以单独的 T 作为终止密码子；柳紫闪蛱蝶 *ATP*6 和 *ND4L* 以 TAG，*ND*3 和 *ND*5 以单独的 T 作为终止密码子（王菊平等，2012；Zhang *et al.*，2012；Chen *et al.*，2012；Wang *et al.*，2013；曹天文等，2013）。从起始密码子和终止密码子的使用情况可以看出，武铠蛱蝶与猫蛱蝶、大紫蛱蝶的同源性较高，黑紫蛱蝶和细带闪蛱蝶居中，与柳紫闪蛱蝶的同源性则较低。同时也再一次验证了昆虫线粒体基因组中 A + T 含量偏向性的特征。

鳞翅目昆虫线粒体基因组 AT 富含区比大多数其他昆虫保守。在 *rrnS* 基因的下游有一段由 ATAGAA 引导的保守的长度为 21 bp 的多聚 T 结构。其余 5 种近缘物种的 AT 富含区也存在着保守的多聚 T 结构，其中柳紫闪蛱蝶中存在的多聚 T 结构长度与武铠蛱蝶相同，为 21 bp，细带闪蛱蝶、猫蛱蝶、大紫蛱蝶、黑紫蛱蝶中存在的多聚 T 结构的长度均为 19 bp。此外，在武铠蛱蝶线粒体 A + T 富含区中存在 2 段长度为 35 bp 的重复单元，其重复序列之间的间隔为 0。并且与其他鳞翅目昆虫比较也存在 $(AT)_8$ 结构。在重复序列上游存在 10 bp 的多聚 A 结构（图 3）。同样地，大紫蛱蝶中存在 2 段长度为 17 bp 的重复单元、1 个 $(AT)_8$ 结构及 1 个 10 bp 的多聚 A 结构；细带闪蛱蝶中存在 2 段长度为 27 bp 的重复单元、1 个 $(AT)_{10}$ 结构及 1 个 9 bp 的多聚 A 结构；猫蛱蝶存在 2 段长度为 15 bp 的重复单元及 1 个 $(AT)_7$ 结构、1 个 $(AT)_8$ 结构及 1 个 9 bp 的多聚 A 结构；黑紫蛱蝶中存在 2 段长度分别为 18 bp 的重复单元和 14 bp 的重复单元及 1 个 $(AT)_8$ 结构；柳紫闪蛱蝶中存在 1 个 $(AT)_7$ 结构、1 个 $(AT)_{10}$ 结构及 1 个 14 bp 的多聚 A 结构。

基于线粒体基因组全序列的 MP 系统发育分析结果中，武铠蛱蝶与其 5 种近缘种的系统进化关系为：(((大紫蛱蝶 *Sasakia charonda* + 黑紫蛱蝶 *Sasakia funebris*) + (细带闪蛱蝶 *Apatura metis* + 柳紫闪蛱蝶 *Apatura ilia*)) + (猫蛱蝶 *Timelaea maculata* + 武铠蛱蝶 *Chitoria ulupi*))。可以看出，武铠蛱蝶 *C. ulupi* 与大紫蛱蝶 *S. charonda*、黑紫蛱蝶 *S. funebris*、细带闪蛱蝶 *A. metis*、柳紫闪蛱蝶 *A. ilia* 和猫蛱蝶 *T. maculata* 的亲缘关系较近，其中，武铠蛱蝶 *C. ulupi* 和猫蛱蝶 *T. maculata* 关系最近。这一结果与传统的形态学分类结果基本一致。

mtDNA 作为一种重要的母性遗传标记被广泛应用于进化研究中，对种群结构和基因流、杂交、生物地理学和系统发育关系的研究都有很重要的意义（Avise *et al.*，1987；Boore，1999；Saccone *et al.*，1999）。本研究对鳞翅目蛱蝶科昆虫武铠蛱蝶线粒体基因组全序列进行了测定并对其线粒体基因组进行了注释和分析，为武铠蛱蝶分子进化和分子鉴定研究提供了科学的依据，同时也为鳞翅目昆虫系统发育提供了基础资料，丰富了鳞翅目昆虫全线粒体基因组研究内容。

参考文献（References）

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC, 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Ann. Rev. Ecol. Syst.*, 18: 489 – 522.

Bae JS, Kim I, Sohn HD, Jin BR, 2004. The mitochondrial genome of the firefly, *Pyrocoelia rufa*: complete DNA sequence, genome organization, and phylogenetic analysis with other insects. *Mol. Phylogenet. Evol.*, 32: 978 – 985.

Ballard JW, 2000. Comparative genomics of mitochondrial DNA in members of the *Drosophila melanogaster* subgroup. *J. Mol. Evol.*, 51: 48 – 63.

Boore JL, 1999. Animal rnitochondrial genomes. *Nucleic Acids Res.*, 27: 1767 – 1780.

Boore JL, 2006. The complete sequence of the mitochondrial genome of *Nautilus macromphalus* (Mollusca: Cephalopoda). *BMC Genomics*, 7: 182.

Boore JL, Lavrov DV, Brown WM, 1998. Gene translocation links insects and crustaceans. *Nature*, 392: 667 – 668.

Buckley TR, Simon C, Chambers GK, 2001. Phylogeography of the New Zealand cicada *Maoricicada campbelli* based on mitochondrial DNA sequences: ancient clades associated with Cenozoic environmental change. *Evolution*, 55(7): 1395 – 1407.

Cameron SL, Lambkin CL, Barker SC, Whiting MF, 2007. A mitochondrial genome phylogeny of Diptera: whole genome sequence data accurately resolve relationships over broad timescales with high

precision. *Syst. Entomol.*, 32: 40 – 59.

Cao TW, Wang JP, Xuan SB, Zhang M, Guo YP, Ma EB, 2013. Analysis of complete mitochondrial genome of *Timelaea maculata* (Lepidoptera, Nymphalidae). *Acta Zootaxonomica Sinica*, 38(3): 468 – 475. [曹天文, 王菊平, 宣善滨, 张敏, 郭亚平, 马恩波, 2013. 猫蛱蝶线粒体基因组全序列分析. 动物分类学报, 38(3): 468 – 475]

Chen M, Tian LL, Shi QH, Cao TW, Hao JS, 2012. The complete mitogenome of the lesser purple emperor *Apatura ilia* (Lepidoptera: Nymphalidae: Apaturinae) and comparison with other nymphalid butterflies. *Zool. Res.*, 33(2): 191 – 201.

Chou I, 1994. Monograph of Chinese Butterflies. Henan Scientific and Technological Publishing House, Zhengzhou. 436 – 437. [周尧 主编, 1994. 中国蝶类志. 郑州: 河南科学技术出版社. 436 – 437]

Chou I, 1998. Classification and Identification of Chinese Butterflies. Henan Scientific and Technological Publishing House, Zhengzhou. 110 – 111. [周尧 主编, 1998 . 中国蝴蝶分类与鉴定. 郑州: 河南科学技术出版社. 110 – 111]

Gene Codes Corporation, 2008. Sequencher, Version 4.8. Gene Codes Corporation, Ann Arbor, Michigan.

Hong GY, Jiang ST, Yu M, Yang Y, Li F, Xue FS, Wei ZJ, 2009. The complete nucleotide sequence of the mitochondrial genome of the cabbage butterfly, *Artogeia melete* (Lepidoptera: Pieridae). *Acta Biochim. Biophys. Sin.*, 41(6): 446 – 455.

Hu J, Zhang D, Hao J, Huang D, Cameron S, Zhu C, 2010. The complete mitochondrial genome of the yellow coaster, *Acraea issoria* (Lepidoptera: Nymphalidae: Heliconiinae: Acraeini): sequence, gene organization and a unique tRNA translocation event. *Mol. Biol. Rep.*, 37(7): 3431 – 3438.

Huang Y, 1998. Molecular Systematics: Principles, Methods and Applications. China Agriculture Press, Beijing. [黄原, 1998. 分子系统学——原理、方法及应用. 北京: 中国农业出版社]

Hwang UW, Park CJ, Yong TS, Kim W, 2001. One-step PCR amplification of complete arthropod mitochondrial genomes. *Mol. Phylogenet. Evol.*, 19: 345 – 352.

Jiang ST, Hong GY, Yu M, Li N, Yang Y, Liu YQ, Wei ZJ, 2009. Characterization of the complete mitochondrial genome of the giant silkworm moth, *Eriogyna pyretorum* (Lepidoptera: Saturniidae). *Int. J. Biol. Sci.*, 5(4): 351 – 365.

Kim MJ, Wan X, Kim KG, Hwang JS, Kim I, 2010. Complete nucleotide sequence and organization of the mitogenome of endangered *Eumenis autonoe* (Lepidoptera: Nymphalidae). *Afr. J. Biotechnol.*, 9(5): 735 – 754.

Kim SR, Kim MI, Hong MY, Kim KY, Kang PD, Hwang JS, Han YS, Jin BR, Kim I, 2009. The complete mitogenome sequence of the Japanese oak silkmoth, *Antheraea yamamai* (Lepidoptera: Saturniidae). *Mol. Biol. Rep.*, 36: 1871 – 1880.

Lee WJ, Kocher TD, 1995. Complete sequence of a sea lamprey (*Petromyzon marinus*) mitochondrial genome: early establishment of the vertebrate genome organization. *Genetics*, 139: 873 – 887.

Liao F, Wang L, Wu S, Li YP, Zhao L, Huang GM, Niu CJ, Liu YQ, Li MG, 2010. The complete mitochondrial genome of the fall webworm, *Hyphantria cunea* (Lepidoptera: Arctiidae). *Int. J. Biol. Sci.*, 6(2): 172 – 186.

Noack K, Zardoya R, Meyer A, 1996. The complete mitochondrial DNA sequence of the bichir (*Polypterus ornatipinnis*), a basal ray-finned fish: ancient establishment of the consensus vertebrate gene order. *Genetics*, 144(3): 1165 – 1180.

Saccone C, De Giorgi C, Gissi C, Pesole G, Reyes A, 1999. Evolutionary genomics in Metazoa: the mitochondrial DNA as a model system. *Gene*, 238(1): 195 – 209.

Singh VK, Mangalam AK, Dwivedi S, Naik S, 1998. Primer premier: program for design of degenerate primers from a protein sequence. *Biotechniques*, 24: 318 – 319.

Tamura K, Dudley J, Nei M, Kumar S, 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Mol. Biol. Evol.*, 24: 1596 – 1599.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG, 1997. The CLUSTALX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.*, 25: 4876 – 4882.

Tzeng CS, Hui CF, Shen SC, Huang PC, 1992. The complete nucleotide sequence of the *Crossostoma lacustre* mitochnodrial genome: conservation and variations among vertebrates. *Nucleic Acids Res.*, 20: 4853 – 4858.

Wang JP, Cao TW, Xuan SB, Wang H, Zhang M, Ma EB, 2013. The complete mitochondrial genome of *Sasakia funebris* (Leech) (Lepidoptera: Nymphalidae) and comparison with other Apaturinae insects. *Gene*, 526(2): 336 – 343.

Wang JP, Nie XP, Cao TW, Zhang M, Guo YP, Ma EB, Zhang XN, 2012. Analysis of complete mitochondrial genome of *Sasakia charonda coreana* (Lepidoptera, Nymphalidae). *Acta Zootaxonomica Sinica*, 37(1): 1 – 9. [王菊平, 聂新平, 曹天文, 张敏, 郭亚平, 马恩波, 张小楠, 2012. 大紫蛱蝶线粒体基因组全序列分析. 动物分类学报, 37(1): 1 – 9]

Wolstenholme DR, 1992. Animal mitochondrial DNA: structure and evolution. *Int. Rev. Cytol.*, 141: 173 – 216.

Xu L, Yu DJ, Zhang RJ, Li JG, Kang L, Chen ZL, Jiao Y, 2007. Sequencing and analysis of the complete mitochondrial genome of the Oriental fruit fly, *Bactrocera dorsalis* (Hendel) (Diptera: Tephritidae). *Acta Entomologica Sinica*, 50(8): 755 – 761. [徐浪, 余道坚, 张润杰, 李建光, 康林, 陈志粦, 焦懿, 2007. 桔小实蝇线粒体基因组全序列及其分析. 昆虫学报, 50(8): 755 – 761]

Zhang F, Mi ZY, 1998. Advance in molecular biology of animal mitochondrial DNA. *Progress in Biotechnology*, 18(3): 25 – 31. [张方, 米志勇, 1998. 动物线粒体 DNA 的分子生物学研究进展. 生物工程进展, 18(3): 25 – 31]

Zhang M, Nie XP, Cao TW, Wang JP, Li T, Zhang XN, Guo YP, Ma EB, Zhong Y, 2012. The complete mitochondrial genome of the butterfly *Apatura metis* (Lepidoptera: Nymphalidae). *Mol. Biol. Rep.*, 39: 6529 – 6536.

Zhang XY, Gao YN, 2004. To design PCR primers with Oligo 6 and Primer Premier 5. *China Journal of Bioinformatics*, 2(4): 15 – 18. [张新宇, 高燕宁, 2004. PCR 引物设计软件使用技巧. 生物信息学, 2(4): 15 – 18]

(责任编辑: 袁德成)